

Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/140955/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Liberatore, Federico ORCID: <https://orcid.org/0000-0001-9900-5108>, Camacho-Collados, Miguel and Quijano-Sánchez, Lara 2022. Equity in the Police Districting Problem: balancing territorial and racial fairness in patrolling operations. *Journal of Quantitative Criminology* 38 , pp. 1-25. 10.1007/s10940-021-09512-x file

Publishers page: <http://dx.doi.org/10.1007/s10940-021-09512-x>
<<http://dx.doi.org/10.1007/s10940-021-09512-x>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.





Equity in the Police Districting Problem: Balancing Territorial and Racial Fairness in Patrolling Operations

Federico Liberatore^{1,2} · Miguel Camacho-Collados^{3,4} · Lara Quijano-Sánchez^{2,5}

Accepted: 2 April 2021
© The Author(s) 2021

Abstract

Objectives The Police Districting Problem concerns the definition of patrol districts that distribute police resources in a territory in such a way that high-risk areas receive more patrolling time than low-risk areas, according to a principle of territorial fairness. This results in patrolling configurations that are efficient and effective at controlling crime but that, at the same time, might exacerbate racial disparity in police stops and arrests. In this paper, an Equitable Police Districting Problem that combines crime-reduction effectiveness with racial fairness is proposed. The capability of this model in designing patrolling configurations that find a balance between territorial and racial fairness is assessed. Also, the trade-off between these two criteria is analyzed.

Methods The Equitable Police Districting Problem is defined as a mixed-integer program. The objective function is formulated using Compromise Programming and Goal Programming. The model is validated on a real-world case study on the Central District of Madrid, Spain, and its solutions are compared to standard patrolling configurations currently used by the police.

Results A trade-off between racial fairness and crime control is detected. However, the experiments show that including the proposed racial criterion in the optimization of patrol districts greatly improves racial fairness with limited detriment to the policing effectiveness. Also, the model produces solutions that dominate the patrolling configurations currently in use by the police.

Conclusions The results show that the model successfully provides a quantitative evaluation of the trade-off between the criteria and is capable of defining patrolling configurations that are efficient in terms of both racial and territorial fairness.

Keywords Police Districting Problem · Smart Patrolling · Racial Inequality · Goal Programming · Compromise Programming

The research of Liberatore is partially funded by the European Commission's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie, grant number MSCA-RISE 691161 (GEO-SAFE), and the Government of Spain, grant MTM2015-65803-R. The research of Quijano-Sánchez was conducted with financial support from the Spanish Ministry of Science and Innovation, grant PID2019-108965GB-I00. All the financial support is gratefully acknowledged.

✉ Federico Liberatore
liberatoref@cardiff.ac.uk

Extended author information available on the last page of the article

Abbreviations

PDP	Police Districting Problem
MIP	Mixed Integer Programming
ST	Spanning Tree
CP	Compromise Programming
EquPDP	Equitable Police Districting Problem
SNPC	Spanish National Police Corps
OECD	Organisation for Economic Co-operation and Development

Introduction

Predictive policing (Perry 2013) consists in the use of analytical quantitative techniques to identify potential criminal activity. Its most popular application is predictive hot spot policing, that is, the allocation of police resources to geographical points which are experiencing an increasing criminal trend, according to predictive models. Optimization models can be used in concert with predictive models to support decision-makers in distributing police resources based on risk maps. In particular, the Police Districting Problem (PDP) (Liberatore et al. 2020) defines efficient and effective patrol sectors according to the distribution of the crime risk in the territory considered. Specifically, PDP produces configurations where the total risk is evenly distributed between the sectors, thus, ensuring that each *r*-district¹ receives an amount of patrolling time that is proportional to its risk-level.

Targeting high crime areas has proven to be successful at reducing crime (Braga et al. 2014). Another consequence of having areas with a higher police presence is that more individuals will be stopped and, eventually, arrested in those areas. It is a well-known fact that high crime locations tend to be correlated with minority and disadvantaged populations (Sampson and Wilson 1995). Therefore, allocating more police presence at hot spots of crime or high crime risk areas has the likely consequence of contributing to racial disparity in police stops and arrests (Andrejevic and Gates 2014; Fagan 2017; Rinehart Kochel 2011). Also, the heightened exposition of minorities to police increases the feeling of pressure in these communities, subsequently producing negative repercussions for both the individuals and society as a whole. On the other hand, reducing police presence in the highest crime areas is not viable, as depriving minority communities of police protection would be discriminatory and could lead to further victimization (Mohler et al. 2015).

In this paper, a model that balances territorial and racial fairness in patrolling operations is proposed, building upon previous research from the authors (Camacho-Collados and Liberatore 2015; Camacho-Collados et al. 2015; Liberatore and Camacho-Collados 2016). The model considers a territory divided into *r*-districts and a population segmented into groups (e.g., racial groups). Ideally, an efficient patrolling configuration assigns each *r*-district an amount of patrolling time that is proportional to its crime risk. At the same time, racial equity is achieved only if each population group has a police contact that is proportional to its size. Therefore, two criteria are maximized: the minimum proportion of service received across all *r*-districts and the minimum proportion of police exposure across all the population groups. Maximizing the minimum proportions increases the patrolling time

¹ *R*-districts (or reporting districts) are the smallest geographical unit for which police statistics are kept and can be considered the atomic territorial element. They can coincide with census blocks.

received by each r -district and population group while reducing the variability as much as possible, thus ensuring efficiency and fairness. Given the distribution of the crime risk and the population on the territory, the model assigns r -districts to patrol sectors and optimizes both criteria according to a weight coefficient which expresses the decision-maker preference for one criterion over the other. By changing its value it is possible to generate the set of efficient solutions which comprise the Pareto frontier.²

The proposed model presents some assumptions and limitations. First, policing involves a large number of different operations; this paper focuses on proactive patrolling operations in a territory. Second, the model does not explicitly incorporate the crimes' severity. On the other hand, the model makes use of the crime risk, which is a versatile measure that can be used to model the stakeholders' preferences as well as the priorities of the police department. For this reason, guidelines on how the crime risk is set are intentionally not provided. However, it could be easily defined to summarise the total level of crime in an area or to implement hot-spot policing. Finally, the proposed approach does not consider game-theoretical elements (Fu and Wolpin 2018; Galiani et al. 2018; Maheshri and Mastrobuoni 2019). As the aim of this paper is to explore the impact of including racial fairness in patrolling decisions, this is left as future research.

The contributions of this article are many-fold. A review of the latest papers on PDP is presented. This review expands and complements the literature review on the subject by Liberatore et al. (2020). The main contribution of this paper lies in the optimization model. To the best of the author's knowledge, this is the first PDP that explicitly considers racial fairness. The model is applied to a real-world case study which allows drawing insights on the usefulness and applicability of the model. The results show that in the dataset considered, a small loss in racial fairness leads to near-optimal solutions in terms of service level. More in general, the analysis of the trade-off between racial fairness and service level is necessary to inform decision-makers and to help them identify solutions that are socially equitable and highly performant.

The rest of the paper is organized as follows. In the next section (Sect. 2) the relevant literature is reviewed. The model is introduced in detail in Sect. 3. Next, in Sect. 4, the case study is introduced and the results of the model are discussed and analyzed. The paper concludes in Sect. 5 with a summary of the findings and a discussion of their implications.

Literature Review

Territory districting (or design) problems are a subfield of discrete optimization related to partitioning decisions (Ríos-Mercado 2020). In its most generic definition, a set of atomic geographic units must be divided into districts or territories, according to specific requirements which depend on the application context. This family of problems has been applied to a wide number of fields, including politics (Kim and Kim 2020; Ricca and Scozzari 2020), health-care (Yanik and Bozkaya 2020; Enayati et al. 2020), sales (Moya-García and Salazar-Aguilar 2020) and, of course, policing.

Since the seminal papers by Mitchell (1972) and Bodily (1978), several police districting models have been presented in the literature. The interested reader is referred to

² The Pareto frontier is the set of all Pareto efficient solutions. A solution is Pareto efficient if no preference criterion can be improved without making at least one preference criterion worse.

the extensive literature review and annotated bibliography on the subject by Liberatore et al. (2020), which is complemented in the rest of this section by presenting alternative approaches to the problem of police patrolling.

Leigh et al. (2019) combine predictive policing and optimization models to target high-crime areas. Hot spots are identified using kernel density estimation, while police agents' positions are determined by a maximum coverage location problems. Another line of research concerns the definition of patrol routes for police agents on a street map. For a review on the subject, the interested reader is referred to Dewinter et al. (2020). Chen et al. (2017) propose a heuristic online Bayesian ant-colony optimization algorithm which generates routes that are efficient, flexible, unpredictable, scalable, and robust. In a subsequent paper (Chen et al. 2019), the authors present a street network police districting problem where the territorial atomic units are street segments rather than r-districts. The objective is to generate patrol sectors that are balanced in terms of workload, which is defined as a combination of multiple factors, i.e., risk, area, and diameter, following previous contributions in the literature (Liberatore et al. 2020). The problem is formulated as a Mixed Integer Programming (MIP) problem and a tabu search meta-heuristic algorithm (Glover 1989, 1990) is proposed for its solution. In their paper, Chen et al. recognize that the street map approach, although it leads to a more straightforward definition of the police patrol routes, is an obstacle to the incorporation of census data in the model, as street segments are incompatible with census districts. A possible solution to this issue is given by Kim (2018) that proposes and tests methods for re-distribute census data to street segments.

To the best of the authors' knowledge, the only optimization model in the context of police patrolling that considers issues of racial inequalities is proposed by Wheeler (2019a). The author tackles the problem of police resources allocation in a fixed districting configuration. The problem is formulated as a Linear Programming model which includes a constraint that imposes an upper bound to the expected police contact of minorities. In another paper, Wheeler (2019b) formulates a police districting model based on the p-median problem. Additional constraints have been added by the author to ensure that the clusters are connected. However, these constraints are incorrect as they still allow a sector to have non-connected subclusters, i.e., the model allows a sector to be formed by geographically disconnected areas. On the other hand, as the focus of the model is on answering calls rather than on proactive patrolling, this does not hinder its usefulness and applicability.

Model

The problem considers a territory partitioned into non-overlapping areas, such as report or census districts. The population in the territory is divided into population groups, that could be defined according to a characteristic of interest, e.g., race, ethnicity or origin. The crime distribution and the population groups distribution are known and, therefore, each area is characterized by a crime level and by the size of each population group. The main decision involves assigning the areas to the patrol districts (multiple assignments are allowed). Then, the time capacity available at each patrol district must be distributed between the tasks of travelling between the areas and patrolling the areas in the district. The patrolling time spent in the areas is used to compute two scores: the service score is the proportion of demand received by an area as patrolling time, while the contact score is the ratio between the police contact time of a population group and its goal. Territorial and racial fairness are achieved by maximizing the minimum service score and the minimum

Table 1 Summary of notation: sets, indices, parameters, and variables

Sets	
$G = (N, E)$	Graph
N	Set of areas
E	Set of edges connecting the areas
K	Number of districts
P	Set of population groups
Indices	
$i \in N$	Area
$k = 1, \dots, K$	District
$p \in P$	Population group
Parameters	
cap	Patrolling time capacity for each district
d_{ip}	Population distribution. Percentage of population in area i that belongs to group p
dem_i	Service demand in area i
$goal_p$	Police contact goal for population group p .
Variables	
$assign_{ik}$	$\begin{cases} 1 & \text{if area } i \text{ is assigned to district } k, \\ 0 & \text{otherwise.} \end{cases}$
$time_{ik}$	Time allocated to area i in district k
ST_k	Length of the Spanning Tree corresponding to district k .
$service_i$	Service level received by area i
W_s	Territorial fairness criterion value.
$contact_p$	Police contact of population group p
W_c	Racial fairness criterion value

contact score, respectively. In fact, Maximizing the minimum score produces solutions with a high average value and low variance.

The rest of the section provides technical details of the methodology. Readers only interested in the practical outcomes of this work may omit this section.

A summary of the notation used in the model is provided in Table 1. The parameters of the model are presented in the following. Let $G = (N, E)$ be an undirected graph where N is the set of areas and E the set of edges connecting them, and let K be the number of districts (indexed by k) and P the set of population groups (indexed by p). Each district has a time capacity, cap , and each population group has a normalized police contact goal (e.g., proportional to the size of the group), $goal_p$, such that $\sum_{p \in P} goal_p = K \cdot cap$. Each area $i \in N$ is characterized by:

- a population distribution d_{ip} representing the percentage of the population of area i that belongs to group p , being $\sum_{p \in P} d_{ip} = 1, \forall i \in N$.
- a normalized service demand (e.g., proportional to the expected crime risk in the area), dem_i , such that $\sum_{i \in N} dem_i = K \cdot cap$.

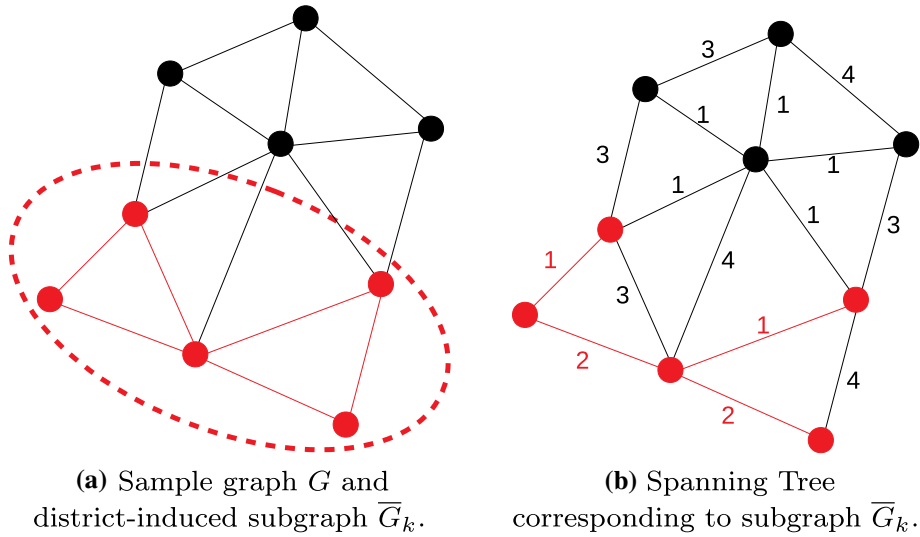


Fig. 1 Sample graph G , **a** district-induced subgraph \bar{G}_k , and **b** corresponding Spanning Tree. **a** The subgraph \bar{G}_k is represented in red. **b** The number next to each edge represent its distance and the edges comprising the ST are drawn in red. The total distance (i.e., the sum of the distances of the edges in ST) is minimum. Also, it is possible to connect any pair of nodes in \bar{G}_k using only edges in ST (Color figure online)

Finally, the edges have an associated distance, $dist_{ij}$, which represents the time required to go from area i to j .

In the problem considered, the main decisions are represented in the model by the following variables:

- $assign_{ik} = \begin{cases} 1 & \text{if area } i \text{ is assigned to district } k \\ 0 & \text{otherwise} \end{cases}$
- $time_{ik} \geq 0$, time allocated to area i by district k .

Each district k defines a subgraph $\bar{G}_k \subset G$ generated by the areas in the district and their incident edges. This is illustrated through an example in Fig. 1a. In the figure, the circled portion of the graph corresponds to a district; the areas that comprise it and the edges that connect them (represented in red) identify a subset, \bar{G}_k , of the original graph G . It is important to notice that exactly defining the route that a patrol would use is out of scope and, also, it is not recommendable as it would make the patrols predictable. Therefore, inside of each district k , the patrol route is approximated by a Spanning Tree (ST) of the graph \bar{G}_k . ST is a subgraph which connects all of the vertices of the original graph at minimum cost. That is, ST identifies the set of edges having minimum total distance that connects all the areas in the district (see Fig. 1b for an example). Implementing ST has the additional advantage of ensuring that the districts are connected, i.e., it is possible to connect every pair of areas in the district using only edges that are part of the district. The interested reader can find the complete formulation of ST in Appendix A.

Let $ST_k \geq 0$ be the length of ST corresponding to district k . It is assumed that a complete patrol route would traverse ST twice as to complete a round trip (i.e., a loop that visits each area at least once and ends where the patrol started) each edge needs to be travelled

two times. Therefore, the capacity of the district is reduced by twice the length of the corresponding ST (i.e., ST_k) and the remaining time can be allocated to its areas. In the following, the constraints and objectives of the model are presented according to their scope.

Districting Constraints

$$\sum_{i \in N} assign_{ik} \geq 1 \quad \forall k = 1, \dots, K \quad (1)$$

$$\sum_{k=1}^K assign_{ik} \geq 1 \quad \forall i \in N \quad (2)$$

$$time_{ik} \leq cap \cdot assign_{ik} \quad \forall i \in N, k = 1, \dots, K \quad (3)$$

$$\sum_{i \in N} time_{ik} = cap - 2 \cdot ST_k \quad \forall k = 1, \dots, K \quad (4)$$

$$assign_{ik} \in \{0, 1\} \quad \forall i \in N, k = 1, \dots, K \quad (5)$$

$$time_{ik} \geq 0 \quad \forall i \in N, k = 1, \dots, K \quad (6)$$

These constraints concern the definition of feasible districts and the allocation of the service time to the areas. The first two sets of constraints (1 and 2) state that no district can be empty and all areas must be assigned to at least one district, respectively. The allocation constraints (3) state that an area can receive patrol time only from districts to which it is assigned. Constraints (4) specify that, for each district, the sum of the patrol time allocated to the areas cannot exceed the district capacity reduced by twice the length of the corresponding ST. The last two sets of constraints present the existence conditions for variables $assign_{ik}$ and $time_{ik}$.

Below, the criteria that allow to explore the trade-off between territorial and racial fairness are presented.

Optimization Criteria

Let us consider an area $i \in N$ and its demand, dem_i . The service level received by an area can be computed as the percentage of demand satisfied, i.e., the ratio between the patrolling time assigned to the area and its demand:

$$service_i = \frac{\sum_{k=1}^K time_{ik}}{dem_i}, \quad \forall i \in N \quad (7)$$

In terms of patrolling efficiency, it is desirable that all the areas receive the highest level of service with little variability among them, to patrol the areas as homogeneously as possible according to their demands and the available capacity. Thus, the territorial fairness criterion is formulated as:

$$\max \min_{i \in N} \{service_i\} \quad (8)$$

which is linearized by introducing an additional continuous variable, W_s , representing the lowest service level across all the areas:

$$\max W_s \quad (9)$$

$$s.t. W_s \leq service_i \quad \forall i \in N \quad (10)$$

On the other hand, the police contact of a population group $p \in P$ can be computed as the ratio between the patrolling time received by the population type and the contact goal:

$$contact_p = \frac{\sum_{k=1}^K \sum_{p \in P} d_{ip} \cdot time_{ik}}{goal_p}, \quad \forall p \in P \quad (11)$$

The equation assumes that all the citizens in an area receive an identical amount of patrolling time, i.e., police contact is assumed to follow a uniform probability distribution. Therefore, the expected police contact of a population group is the sum across the areas of the proportion of time received by the area relative to the size of the group.

Police contact should be as homogeneous as possible among the groups to ensure racial fairness. At the same time, it is still important to provide each group with the highest possible level of police contact, to avoid social abandonment and victimization from crime. Therefore, the racial fairness criterion can be formulated as:

$$\max \min_{p \in P} \{contact_p\} \quad (12)$$

which is linearized by introducing an additional continuous variable, W_c , representing the lowest exposition level across all the groups:

$$\max W_c \quad (13)$$

$$s.t. W_c \leq contact_p \quad \forall p \in P \quad (14)$$

Payoff Matrix and Compromise Programming Model

To verify the trade-off between the criteria, the payoff matrix is computed:

$$\begin{pmatrix} W_s^+ & W_c^- \\ W_s^- & W_c^+ \end{pmatrix} \quad (15)$$

The matrix presents for each criterion (i.e., W_s and W_c) the ideal and the anti-ideal values, identified by a '+' and a '-' superscript, respectively. The ideal and the anti-ideal values are the best and worst possible values obtainable for each criterion, respectively. The payoff matrix is computed following Algorithm 1.

Algorithm 1 Payoff matrix algorithm.

- | | |
|--|-----------------------------------|
| 1: $W_s^+ \leftarrow \text{optimize}\{(9), (10), (7), (1) - (6), (23) - (35)\}$ | ▷ Territorial fairness ideal |
| 2: $W_c^- \leftarrow \text{optimize}\{(13), (14), (11), (10), (7), W_s \geq W_s^+, (1) - (6), (23) - (35)\}$ | ▷ Racial fairness anti-ideal |
| 3: $W_c^+ \leftarrow \text{optimize}\{(13), (14), (11), (1) - (6), (23) - (35)\}$ | ▷ Racial fairness ideal |
| 4: $W_s^- \leftarrow \text{optimize}\{(9), (10), (7), (14), (11), W_c \geq W_c^+, (1) - (6), (23) - (35)\}$ | ▷ Territorial fairness anti-ideal |

Each step of the algorithm computes one of the values in the payoff matrix (15). First, the ideal territorial fairness (W_s^+) is calculated by maximizing the minimum service level across the districts. Then, the anti-ideal racial fairness (W_c^-) is identified by maximizing the minimum police contact while imposing that the territorial fairness must be at least as high as its ideal. This completes the first row of the payoff matrix. Next, the ideal racial fairness (W_c^+) is determined by maximizing the minimum police contact and, finally, the anti-ideal territorial fairness (W_s^-) is assessed by maximizing the minimum service level across the districts while imposing that the racial fairness must be at least as high as its ideal. After this step, the full payoff matrix is achieved.

To obtain a solution to this multi-criteria problem, Compromise Programming (CP) (Ringuest 1992) is used. CP is a decision making technique that finds the best solution by minimizing the linear combination of the normalized distance of the criteria values to the ideal point. In this model the distance value is calculated as:

$$Z = \lambda \cdot \frac{W_s^+ - W_s}{W_s^+ - W_s^-} + (1 - \lambda) \cdot \frac{W_c^+ - W_c}{W_c^+ - W_c^-} \quad (16)$$

where $\lambda \in [0, 1]$ is a parameter that specifies the preference of the decision-maker on the territorial fairness criterion over the racial fairness criterion. The complete model, called the Equitable Police Districting Problem (EquPDP), is given below.

[EquPDP]

min Z

Compromise programming distance: (16)

Optimization criteria: (7), (10), (11), (14)

Partitioning constraints: (1)–(6)

Spanning tree constraints: (23)–(35)

Let K be the number of districts to be defined, $|N|$ the number of areas, $|P|$ the number of population groups, and $|E|$ the number of edges. Then, the model includes: $(2K|N| + K|E|)$ binary variables, $(3 + K + |P| + |N| + 2K|N| + K|E|)$ real variables, and $(1 + 4K + 2|P| + 3|N| + 9K|N| + 5K|E|)$ constraints (including the existence conditions for the variables). The solutions provided by EquPDP are efficient and by varying the value of λ it is possible to obtain the Pareto Frontier.

Case Study

EquPDP is applied to a real case study on the Central District of Madrid, Spain. The Central District of Madrid is approximately 5.23 km² (2.02 sq mi) in size, it has a population of 149,718 people and a population density of 28,587/km² (74,040/sq mi). Administratively,

the district is divided into six wards (also called *barrios*): *Cortes*, *Embajadores*, *Justicia*, *Universidad*, *Palacio*, and *Sol*.

The crime risk definition given below has been determined according to the objectives of the Spanish National Police Corps (SNPC) and the Central District of Madrid PD in particular. According to the OECD Better Life Index, Spain actually has a lower incidence of crime than in many other OECD countries (OECD 2020), e.g., Spain has a homicide rate of 0.6 murders per 100,000 inhabitants, compared to the OECD average of 3.6 per 100,000. In particular, the capital city of Madrid and its surrounding region is the safest of the large population centers; in fact, the region of Madrid contains 13.75% of Spanish residents but it is responsible for just 10.41% of all crime. Theft is the most frequent type of crime committed in Spain and one of the main priorities for the SNPC is its reduction. For the above reasons, in this case study the crime risk of an area is the number of thefts reported.

Dataset

The topological data has been obtained from official sources (NOMECALLES 2020). Census districts have been chosen as the atomic territorial units. Each census district is represented by its centroid, which translates into an area in the graph on which the model is formulated. Two areas are connected by an edge if the corresponding census districts share part of their perimeters. The territory and the resulting graph, comprised of 111 areas and 308 edges, are shown in Fig. 2.

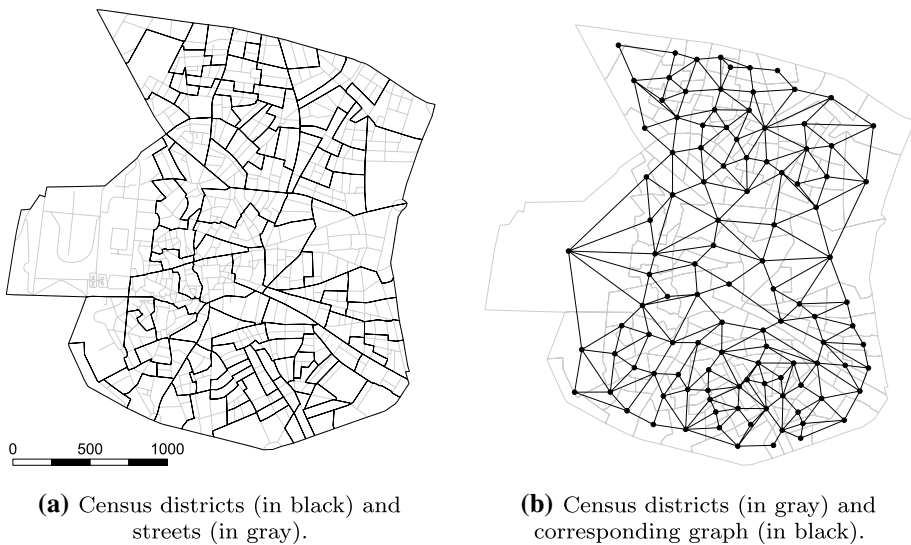


Fig. 2 Central District of Madrid. Graphical representation of streets, census districts, and corresponding graph. The scale is in meters (Color figure online)

For each edge $(i, j) \in E$, let $length_{ij}$ be the great-circle distance³ between the areas in meters. Then, the edges' distance is calculated as follows:

$$dist_{ij} = length_{ij}/100 \quad \forall (i, j) \in E \quad (17)$$

which corresponds to the time in minutes that it takes to walk the edge assuming a walking speed of 100 m/min (equivalent to 6 km/h).

The crime data has been provided by the SNPC. Specifically, the dataset includes all the reported thefts occurred in each area during the following shifts:

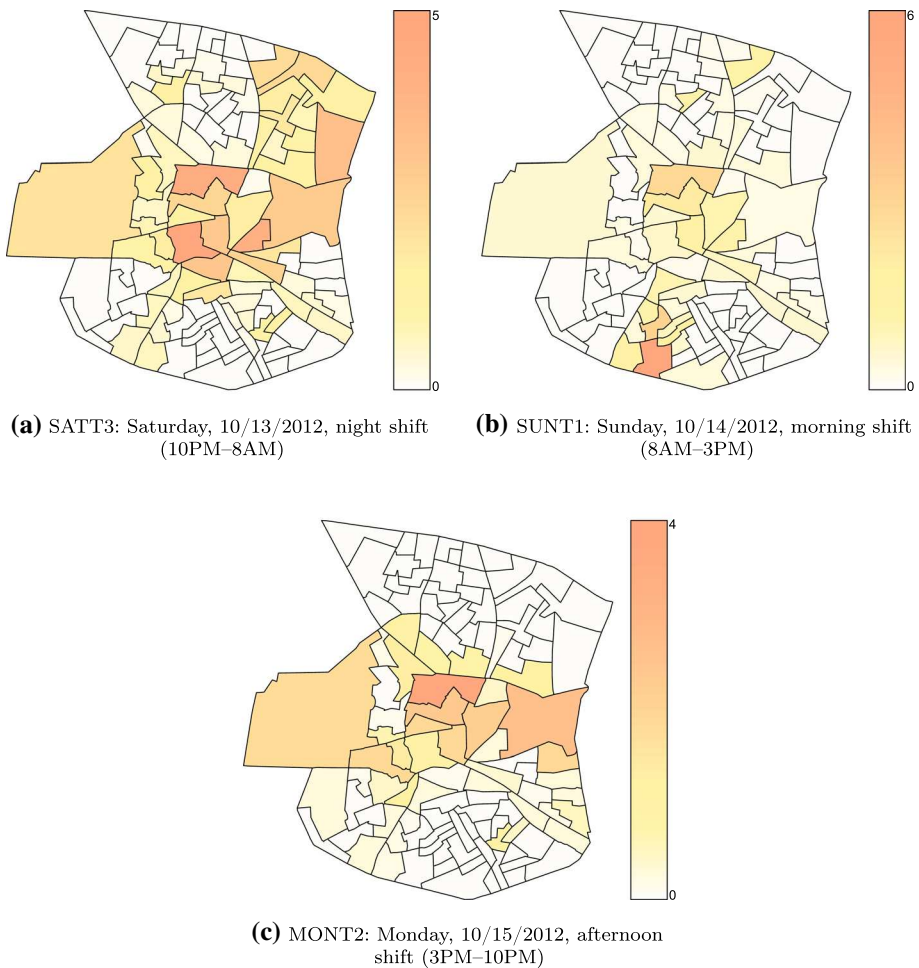


Fig. 3 Heat maps for the Central District of Madrid corresponding to three different shifts. The census districts' color identify the associated crime level, ranging from red for high crime level to white for no criminal activity, as specified in the color legend (Color figure online)

³ The great-circle distance is the shortest distance between two points on the surface of a sphere, measured along the surface of the sphere. It is computed as the length of the arc linking the two points on the sphere. This provides a better approximation of the distance between two points on Earth's surface than the Euclidean's distance.

- SATT3: Saturday, 10/13/2012, night shift (10PM–8AM).
- SUNT1: Sunday, 10/14/2012, morning shift (8AM–3PM).
- MONT2: Monday, 10/15/2012, afternoon shift (3PM–10PM).

The heat maps for each shift are represented in Fig. 3.

Let $risk_i$ be the number of thefts in area i . Then, the areas' demand is set to

$$dem_i = \frac{risk_i}{\sum_{i' \in N} risk_{i'}} \cdot K \cdot cap \quad \forall i \in N \quad (18)$$

where cap is the duration of the shift in minutes, i.e., 600 for SATT3 and 420 for SUNT1 and MONT2. Concerning the number of districts, the values considered are $K = \{2, 6\}$. In fact, the standard patrolling configurations for the Central District of Madrid adopted by SNPC (represented in Fig. 4) partition the territory into either two districts (i.e., north/south of the main artery, the *Gran Vía*) or six (i.e., according to the wards), which are each assigned to multiple officers.

The population data has been obtained from the 2011 Spanish Census (INE 2011) which provides data on a census district level and segments the population according to the following geographical regions of birth: (a) Spain; (b) other EU country; (c) European non-EU country; (d) Africa; (e) Caribbean, Central and South America; (f) North America; (g) Asia; (h) Oceania. Following this categorization, the population groups set P has eight elements, one for each region of birth. The distribution of each population group is illustrated in Fig. 5.

Let pop_{ip} be the number of people living in i that were born in region p . Then, the population distribution over the territory and the patrolling goals for the population groups are calculated as follows:

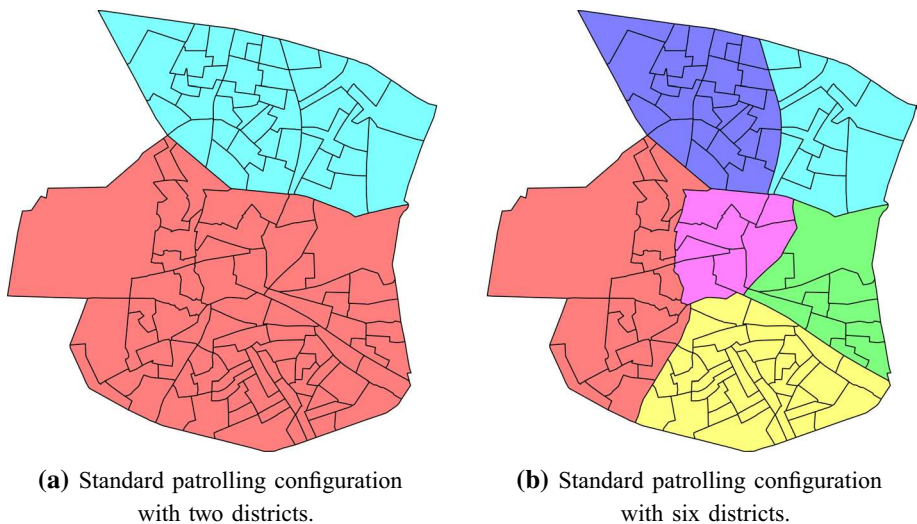


Fig. 4 Standard patrolling sectors in the Central District of Madrid. The census districts' borders are plotted in black and each patrol sector is represented using a different color (Color figure online)

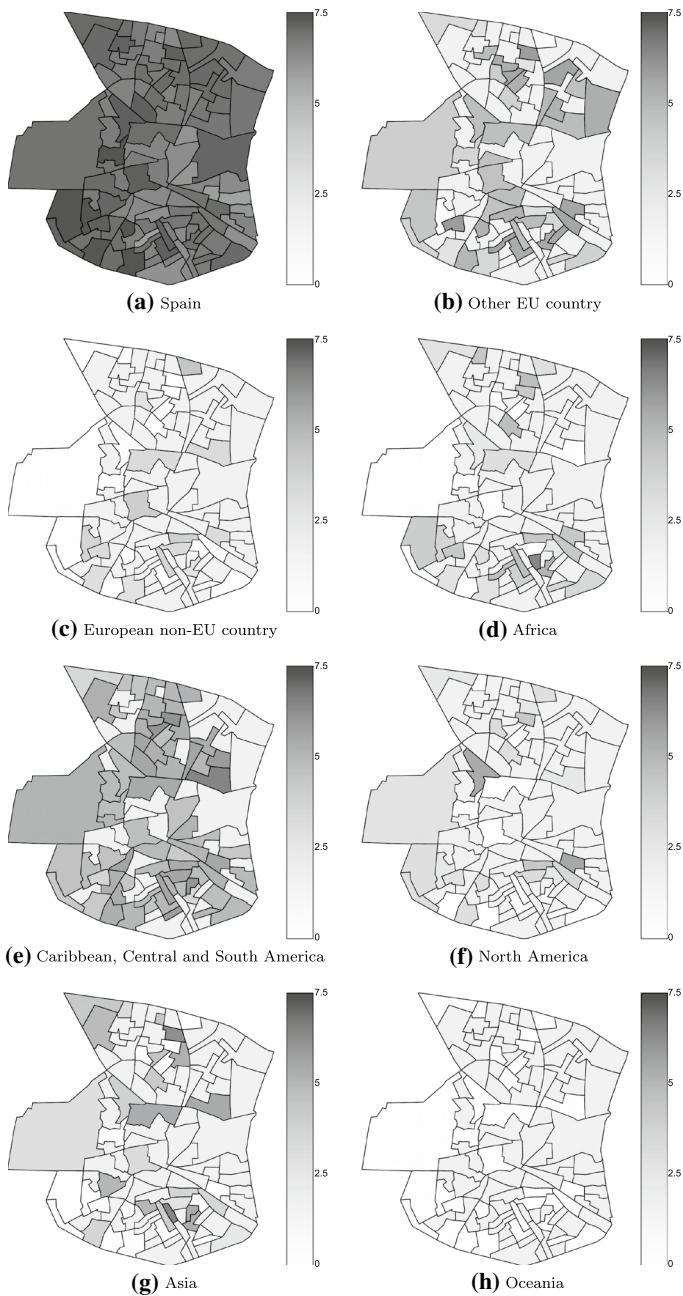


Fig. 5 Population distribution on the territory by population group. The number of inhabitants is represented using a gray palette, ranging from dark gray for high values to white for no inhabitants, according to the color legend. Due to the predominance of the first group (i.e., Spain) over the others, the values are represented on the logarithmic scale (Color figure online)

$$d_{ip} = \frac{pop_{ip}}{\sum_{p' \in P} pop_{ip'}} \quad \forall i \in N, p \in P \quad (19)$$

$$goal_p = \frac{\sum_{i \in N} pop_{ip}}{\sum_{i \in N} \sum_{p' \in P} pop_{ip'}} \cdot K \cdot cap \quad \forall p \in P \quad (20)$$

Computational Experiments, Results and Discussion

In this section, the experiments and their results are presented and discussed. The following experiments are conducted:

- Computation and analysis of the payoff matrices to verify the existence of a trade-off between the criteria.
- Analysis of the police contact levels in the payoff matrices' solutions to verify the effect of the racial fairness criterion (Equation 12).
- Sampling of the Pareto frontiers and comparison with the standard patrolling configuration adopted by SNPC.

The algorithm has been programmed in Julia v.1.4.0 (Bezanson et al. 2017). EquPDP has been implemented in JuMP v.0.21.2 (Dunning et al. 2017) and solved using Gurobi v.9.0.1. (Gurobi Optimization 2020). All the experiments have been run on a Dell Precision 5540, equipped with Intel® Core™ i9-9880H CPU @ 2.30GHz × 16 and 16GB RAM. The standard configuration of Gurobi has been used, which applies multithreading. A CPU time limit of 3600s has been set on all the optimization processes.

Overall, six instances are considered, obtained by combining three shifts (i.e., SATT3, SUNT1, and MONT2) with two values of K (i.e., $K = 2, 6$). For each instance the payoff matrix is computed and the Pareto frontier is sampled at the following values of the parameter $\lambda = \{0.0, 0.25, 0.5, 0.75, 1.0\}$. This entails solving nine optimization models per instance, i.e., four to compute the payoff matrix and one for each value of λ . Due to the size of the graph, none of the optimization processes could complete within the time limit. Therefore, the results of EquPDP presented below are sub-optimal. On the other hand, the evaluation of the standard configurations by SNPC involves solving EquPDP with fixed assignment variables (i.e., $assign_{ik}$) and takes less than one second to complete to optimality.

Payoff Matrices

Table 2 illustrates the payoff matrices for the instances considered for $K = 2$ and $K = 6$. It is possible to identify a trade-off between the criteria. More in detail, when $K = 6$ it can be appreciated that the territorial fairness criterion presents a larger gap between the ideal and the anti-ideal value than the racial fairness criterion. Also, when optimizing with respect to one criterion, the model achieves a high level of efficiency, higher than 0.85 for both criteria. On the other hand, when $K = 2$, the criteria present similar values and gaps between the ideal and the anti-ideal points.

Table 2 Payoff matrices

Shift	K	W_s^+	W_s^-	W_c^+	W_c^-
<i>SATT3</i>	2	0.69946	0.47181	0.70164	0.40491
<i>SUNT1</i>	2	0.56946	0.30542	0.57377	0.36478
<i>MONT2</i>	2	0.57017	0.29924	0.57377	0.31030
<i>SATT3</i>	6	0.89683	0.04805	0.90806	0.51911
<i>SUNT1</i>	6	0.85503	0.09886	0.86826	0.54447
<i>MONT2</i>	6	0.85446	0.01543	0.86863	0.46883

The columns present the following information: shift, number of districts (K), territorial fairness ideal and anti-ideal values (columns W_s^+ and W_s^- , respectively), racial fairness ideal and anti-ideal values (columns W_c^+ and W_c^- , respectively)

Racial Fairness Criterion Analysis

To verify the ability of the racial fairness criterion to generate solutions with homogeneous police contact levels across all the populations groups, the solutions obtained by the steps 2 and 4 of Algorithm 1 are considered and compared. In particular, these steps are equivalent to solving EquPDP with the following objective functions:

$$\text{lex max} \{ W_s, W_c \} \quad (21)$$

and

$$\text{lex max} \{ W_c, W_s \} \quad (22)$$

where “lex” stands for “lexicographic”.⁴ Equation (21) prioritizes W_s over W_c , yielding solutions that are more efficient yet less equitable in terms of population contact with the police, while Equation (22) does the opposite.

Table 3 illustrates the average value and the standard deviations for the population groups’ contact levels with the police on all the problem instances considered. It is evident

Table 3 Comparison of police contact levels in solutions corresponding to objective functions (21) and (22)

Shift	K	obj. fun. (21) avg \pm sd	obj. fun. (22) avg \pm sd
<i>SATT3</i>	2	0.65641 \pm 0.18442	0.70164 \pm 0.0
<i>SUNT1</i>	2	0.52099 \pm 0.12128	0.57377 \pm 0.0
<i>MONT2</i>	2	0.53699 \pm 0.13693	0.57377 \pm 0.0
<i>SATT3</i>	6	0.84128 \pm 0.23668	0.90811 \pm 0.00013
<i>SUNT1</i>	6	0.78192 \pm 0.18262	0.86826 \pm 0.0
<i>MONT2</i>	6	0.80495 \pm 0.20441	0.86863 \pm 0.0

The columns present the following information: shift, number of districts (K), average and standard deviation of contact level (variables $contact_p$) for objective function (21) and (22)

⁴ The lexicographic method assumes that a set of objectives can be ranked in decreasing order of importance. It consists of iteratively solving a sequence of single-objective optimization problems, one per objective. At each iteration, the corresponding objective is optimized, while the values of the higher-ranking objectives are fixed to the values obtained in the previous steps. The solution to the program is given when the last objective is optimized.

that prioritizing the racial fairness criteria results in solutions that have almost no variability and high contact level on average. On the other hand, optimizing with respect to territorial fairness yields solutions that are highly unequal in terms of police contact, with standard deviations ranging from 0.12 to 0.24, approx. Therefore, it can be concluded that the racial fairness criterion indeed produces the desired results.

Pareto Frontiers and Comparison with SNPC Configurations

Figure 6 illustrates the sampled Pareto frontiers for the considered instances and the corresponding SNPC patrolling configurations (see Fig. 4). The figure is comprised of six plots, one per problem instance (i.e., SATT3, SUNT1, and MONT2) and value of $K = 2, 6$. Each plot is a scatterplot of the solution values found by EquPDP and of the SNPC solution at different values of the preference coefficient $\lambda = 0, 0.25, 0.5, 0.75, 1$. The axis correspond to the fairness criteria values (i.e., the territorial fairness, W_s , and the racial fairness, W_c , criteria). In each scatterplot, the solution points are drawn differently depending of whether they are generated by EquPDP (dot) or are the SNPC solution (triangle). Also, the points are connected by lines (continuous and dotted, respectively) to show the corresponding frontier. It can be seen that, in every plot, the frontier corresponding to EquPDP is always on top of the SNPC's frontier; this means that the solutions defined by EquPDP always dominate the SNPC's standard configuration. This means that EquPDP's solutions are always better than the SNPC's standard configuration and, therefore, the proposed model clearly improves on the quality of the patrolling operations currently in use.

On a different note, when considering only the solutions found by EquPDP it can be seen that the frontiers present a very clear elbow. More importantly, the curves show that non-extreme solutions (i.e., the solutions obtained when $0.25 \leq \lambda \leq 0.75$) provide a good compromise between the two criteria. Given the proximity of the non-extreme solution points in the plots, the value of the parameter λ does not seem to have a major impact on the criteria values, as long as it is $0.25 \leq \lambda \leq 0.75$.

It is important to notice that, as mentioned above, the solutions obtained in this experiments by the EquPDP are sub-optimal, which means that their value (variable Z , Eq. 16) is an upper bound to the optimum (Z^*). In a plot where the axes are W_s and W_c (such as in Fig. 6), the point corresponding to a sub-optimal solution would be closer to the origin than the point corresponding to the optimal one. Therefore, the Pareto frontier identified by optimal solutions would be farther away from the origin than the current one and, as a consequence, the dominance of the solutions found by EquPDP over the solution of the SNPC would be even greater. This is a further proof of the superiority of EquPDP over the standard configurations currently in use.

Table 4 compares the criteria values for all instances at $\lambda = 0.0$ and $\lambda = 0.25$. The results displayed in the table confirm that, by sacrificing approximately 0.01 in racial fairness, it is possible to improve the service level by more than 0.50 for $K = 2$, and 0.75 for $K = 6$. This outcome is specific to the dataset considered; in fact, the gains in terms of racial fairness depend on the level of segregation among population groups. In particular, it is expected for these gains to be lower for higher levels of segregation. Regardless of that, what these results manifest is that the trade-off between racial and service equity needs to be assessed to make informed decisions, as high level of racial fairness could still lead to nearly-optimal solutions in terms of service level.

Figures 7 and 8 shows the patrolling configurations corresponding to the solutions produced by EquPDP for $K = 2, 6$ and $\lambda = 0.25$. Despite EquPDP allows to assign an area to

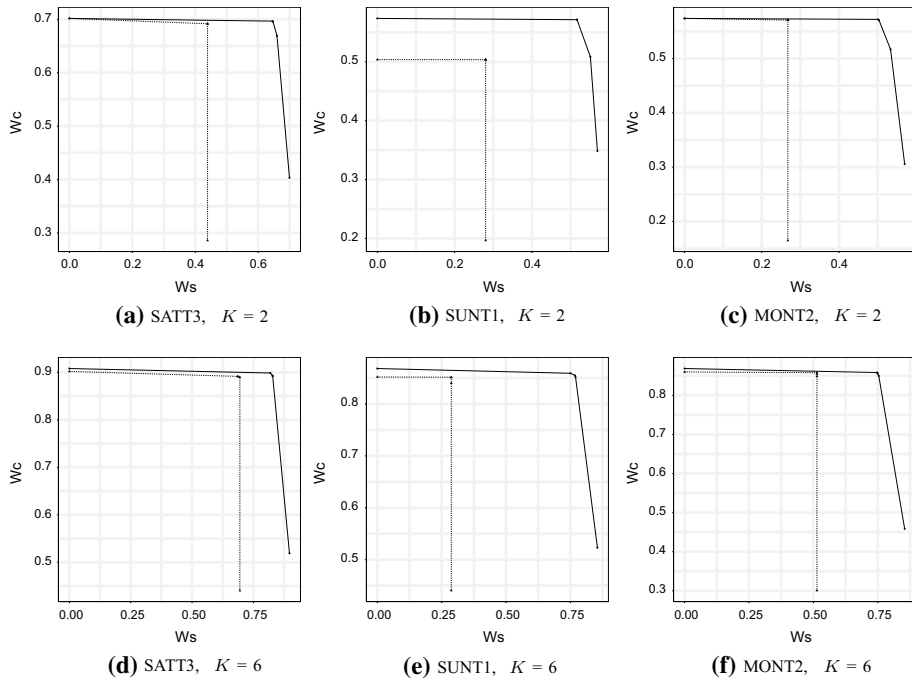


Fig. 6 Comparison of Pareto frontiers for EquPDP (straight line) and SNPC patrolling configurations (dotted line), evaluated at $\lambda = \{0.0, 0.25, 0.5, 1.0\}$

Table 4 Criteria values at $\lambda = 0.0$ and $\lambda = 0.25$

Shift	K	W_s at $\lambda = 0.0$	W_s at $\lambda = 0.25$	W_c at $\lambda = 0.0$	W_c at $\lambda = 0.25$
<i>SATT3</i>	2	0.0	0.64549	0.70164	0.69656
<i>SUNT1</i>	2	0.0	0.51677	0.57377	0.57161
<i>MONT2</i>	2	0.0	0.50066	0.57377	0.57196
<i>SATT3</i>	6	0.0	0.81898	0.90807	0.89861
<i>SUNT1</i>	6	0.0	0.75052	0.86853	0.85912
<i>MONT2</i>	6	0.0	0.74859	0.86863	0.85836

The columns present the following information: shift, number of districts (K), territorial fairness (W_s) at $\lambda = 0.0$ and $\lambda = 0.25$, and racial fairness (W_c) at $\lambda = 0.0$ and $\lambda = 0.25$

multiple district (see Eq. 2) only two configurations have overlapping patrol districts. The solutions for $K = 2$ (Fig. 7) are quite similar, with small variations to adapt to the idiosyncrasies of each shift. On the other hand, for $K = 6$ (Fig. 8) very different configurations are obtained:

- In shift SATT3, most crimes are concentrated in the middle section of the district and in the north-eastern corner (see Fig. 4a), which represent approximately 50% of the ter-

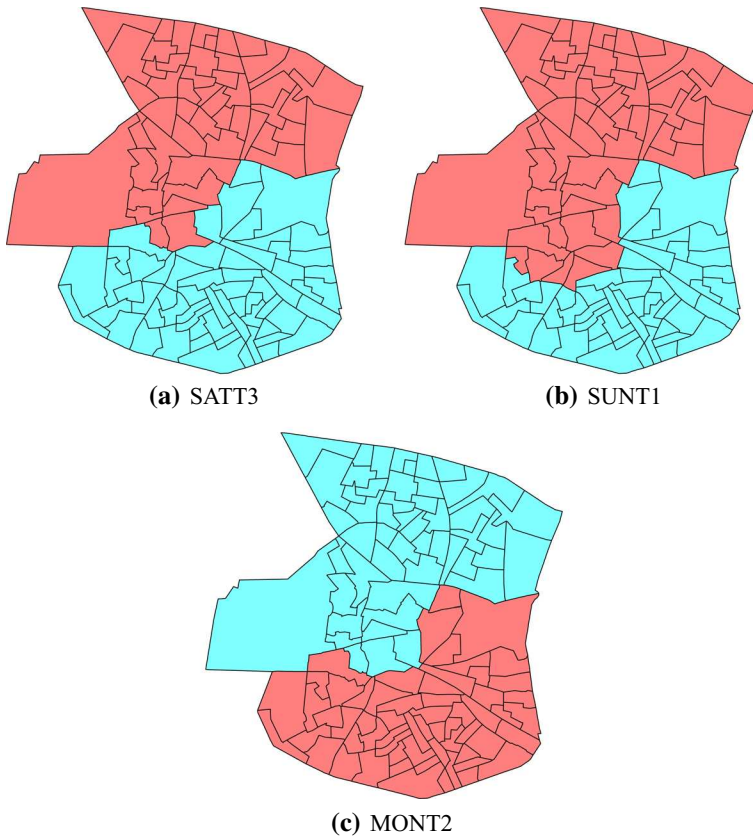


Fig. 7 Patrolling configurations obtained by EquPDP for $K = 2$ and $\lambda = 0.25$. Each patrol sector is represented using a different color (Color figure online)

ritory and correspond to busy nightlife hot spots. In the solution identified by EquPDP this area is covered by as many as five different sectors.

- Concerning shift SUNT1, most crimes are located in the center and in the south. The census district with the highest crime rate corresponds to a flea market that takes place every Sunday morning. In the solution, the yellow sector covers the center while the dark blue sector focuses on the area of the flea market. Three sectors (green, light blue, and red) provide support to cover the south, while the pink sector is dedicated exclusively to the north.
- During shift MONT2, crime is concentrated almost exclusively in the middle section of the district, corresponding to about one third of the territory. EquPDP produces a solution where this area is covered by five different sectors, while the red sector focuses on the north, which is almost entirely free of crime.

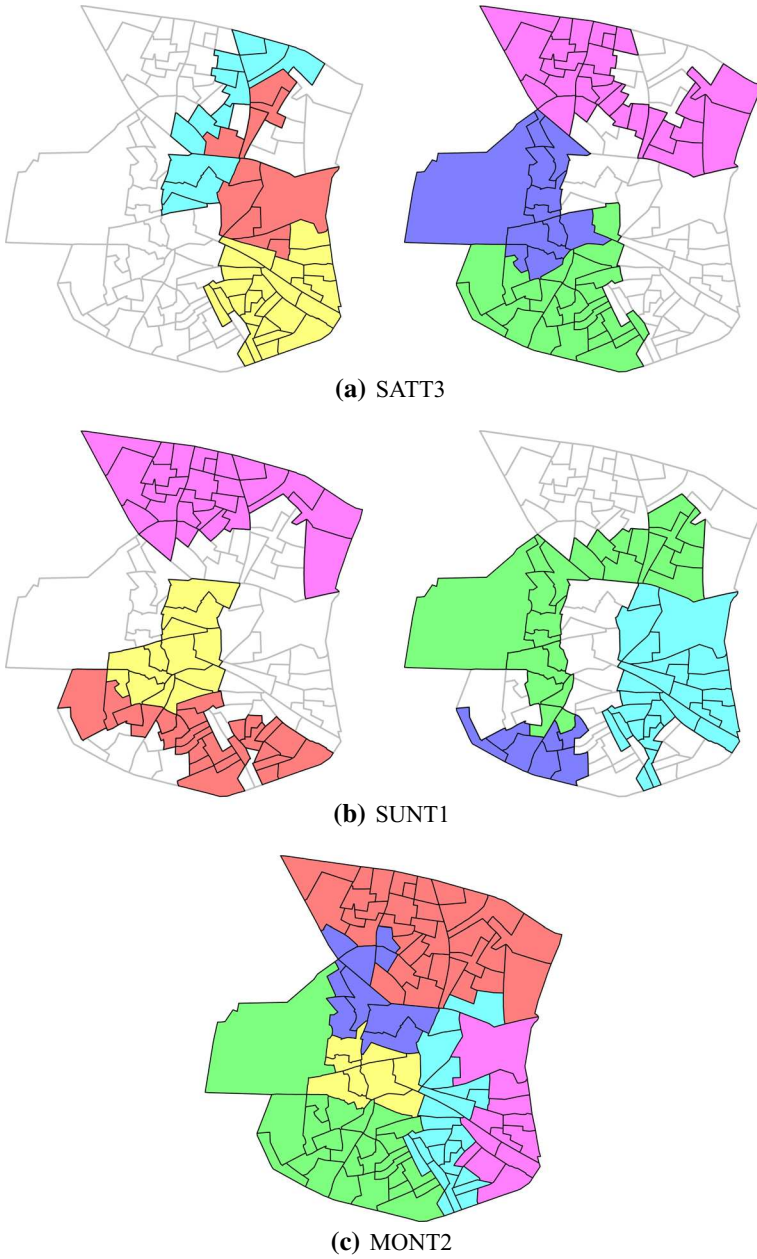


Fig. 8 Patrolling configurations obtained by EquPDP for $K = 6$ and $\lambda = 0.25$. Each patrol sector is represented using a different color. Patrolling configurations having overlapping sectors are represented using multiple maps (Color figure online)

Conclusions

In this paper, a police districting problem that balances between territorial and racial fairness is proposed. The problem is formulated as a MIP model that defines patrolling configurations that are operationally efficient and, at the same time, equitable in terms of the contact between the police force and the different population groups. The model is tested on a real-world case study on the Central District of Madrid, Spain, which allows drawing some insights regarding the model that are presented in the following.

Results Summary

The following highlights the major findings obtained from the experiments.

- There is a clear trade-off between territorial and racial fairness.
- For $K = 6$, the gap between the ideal and the anti-ideal values is larger for the territorial fairness criterion. Also, solutions that prioritize racial fairness result in a service level close to zero. On the other hand, prioritizing territorial fairness produces solutions with a racial fairness close to 0.50. For $K = 2$ these differences are not observed.
- The racial fairness criterion effectively produces solutions with high average police contact level and extremely low variability (i.e., standard deviation close to zero).
- The solutions identified by EquPDP dominate the standard patrolling configurations adopted by SNPC.
- The Pareto frontier presents a well-defined elbow. Furthermore, non-extreme solutions are clustered in proximity of this elbow and are close to the ideal point. This implies that the value of parameter λ does not have a major impact on the criteria values and that, as long as it is strictly larger than zero and smaller than one, the solution found by EquPDP will be “good.”
- In terms of compromise between criteria, a very small decrease in racial fairness (0.01 approx.) results in a great improvement in territorial fairness (higher than 0.70).
- Thus, it is extremely beneficial to optimize considering both criteria at the same time, as it results in solutions that are both very effective at controlling crime and equitable.

These insights are specific to the case study considered; however, they illustrate the applicability and usefulness of the methodology proposed. In particular, the results show that the Pareto frontier has a very sharp elbow. This indicates that most compromise solutions are clustered in the proximity of the ideal point and, therefore, choosing any among them does not have a major impact on performances. Optimizing exclusively for territorial fairness results in patrolling configurations that are highly imbalanced in terms of racial fairness and, vice-versa, prioritizing racial fairness produces very inefficient solutions. On the other hand, by sacrificing approximately 0.01 in racial fairness, it is possible to significantly improve the service level. This demonstrates the importance for the decision-makers of studying the trade-off between racial and territorial fairness.

Discussion

The experimental results verify that, as postulated by Kleinberg et al. (2018), the explicit incorporation of racial information into police decision making can lead to more fair and efficient outcomes. On the other hand, disregarding this information might lead to unfair

results which have the potential to exacerbate racial disparity, due to the structural differences between population groups and the correlation between crime level and minorities presence.

The approach presented in this paper hinges on improving racial fairness at the expenses of effectiveness. Certainly, voluntarily depriving high-crime areas of much-needed security gives rise to an ethical dilemma. Although the methodology introduced allows quantifying the trade-off between effectiveness and racial fairness, it does not provide a solution to such predicament. What is usually recommended in these cases is that the decision-makers (e.g., law enforcement agencies and/or the public) should determine the degree of preference of racial fairness over effectiveness (Cohen 2017), that is represented in the model by the coefficient λ . However, it is the opinion of the authors that the ethical dilemma described above arises only when looking at the crime-reduction benefits of police operations in the short and medium term. Recent events (Wikipedia contributors 2020) have clearly shown the consequences that racial inequality in police interventions can have on society in the long term. These consequences include protest marches, social unrest and riots, which might lead to a surge in violence and crime. Factoring these long-term effects provides sufficient ground to estimate with objectivity and precision the most beneficial level of preference between the two criteria. This would require longitudinal and correlation study, and is left as future research.

The methodology presented in this paper does not solve the problem of racial inequality in police interventions. However, it can be used in conjunction with other strategies intended to mitigate long term harm to particular communities, such as problem-oriented policing (Scott and Clarke 2020), proactive policing (National Academies of Sciences 2018), and increasing the perceived legitimacy of police interactions with the public (Braga et al. 2019).

The authors hope that this work will be a useful source of ideas for future research on PDP and will contribute further in the development and solution of more complex and more realistic models in the context of policing, public security and safety.

Appendix A Spanning Tree Formulation

In the following, the mathematical programming model for the ST is provided. ST is formulated using a flow model for each district. One of the areas in the district is selected as the “root” that emits supply while the other areas must receive one unit of supply each. A flow of supply travels from area to area on the edges; all the non-root areas must satisfy the condition that the amount of entering flow must be equal to the amount of flow that leaves it plus one extra unit of supply that each area must receive. This ensures that the corresponding district is connected, as all the areas in ST must be reachable from the root of ST using only edges in the district (Shirabe 2005; Duque et al. 2011).

A summary of the notation is given in Table 5. The variables required to model ST are:

- $ST_k \geq 0$, length of ST corresponding to district k .
- $root_{ik} = \begin{cases} 1 & \text{if area } i \text{ is the root of ST of district } k \\ 0 & \text{otherwise} \end{cases}$
- $cross_{ijk} = \begin{cases} 1 & \text{if edge } (i, j) \text{ is in ST of district } k \\ 0 & \text{otherwise} \end{cases}$

Table 5 Additional notation for the Spanning Tree model. Expands on the notation in Table 1

Indices	
$j \in N$	Area
$(i, j) \in E$	Edge
Parameters	
$dist_{ij}$	Length of edge (i, j)
Variables	
$root_{ik}$	$\begin{cases} 1 & \text{if area } i \text{ is the root of the Spanning Tree of district } k \\ 0 & \text{otherwise} \end{cases}$
$cross_{ijk}$	$\begin{cases} 1 & \text{if edge } (i, j) \text{ is in the Spanning Tree of district } k \\ 0 & \text{otherwise} \end{cases}$
$supply_{ik}$	Flow supply for area i in district k
$flow_{ijk}$	Flow traversing edge (i, j) in district k

- $supply_{ik} \geq 0$, supply of area i in the flow model of district k . This variable takes value greater than zero only for the root area of ST, i.e., $root_{ik} = 1$. In this case, its value is the number of areas in k .
- $flow_{ijk} \geq 0$, flow on the edge (i, j) in district k . This variable can take value greater than zero only for the edges in ST, i.e., $cross_{ijk} = 1$.

$$ST_k = \sum_{(i,j) \in E} dist_{ij} \cdot cross_{ijk} \quad \forall k = 1, \dots, K \quad (23)$$

$$cross_{ijk} \leq assign_{ik} \quad \forall (i, j) \in E, k = 1, \dots, K \quad (24)$$

$$cross_{ijk} \leq assign_{jk} \quad \forall (i, j) \in E, k = 1, \dots, K \quad (25)$$

$$flow_{ijk} \leq |N| \cdot cross_{ijk} \quad \forall (i, j) \in E, k = 1, \dots, K \quad (26)$$

$$\sum_{i \in N} root_{ik} = 1 \quad \forall k = 1, \dots, K \quad (27)$$

$$root_{ik} \leq assign_{ik} \quad \forall i \in N, k = 1, \dots, K \quad (28)$$

$$supply_{ik} \leq N \cdot root_{ik} \quad \forall i \in N, k = 1, \dots, K \quad (29)$$

$$supply_{ik} \leq \sum_{j \in N} assign_{jk} \quad \forall i \in N, k = 1, \dots, K \quad (30)$$

$$supply_{ik} + \sum_{j: (j,i) \in E} flow_{jik} = \sum_{j: (i,j) \in E} flow_{ijk} + assign_{ik} \quad \forall i \in N, k = 1, \dots, K \quad (31)$$

$$root_{ik} \in \{0, 1\} \quad \forall i \in N, k = 1, \dots, K \quad (32)$$

$$cross_{ijk} \in \{0, 1\} \quad \forall (i, j) \in E, k = 1, \dots, K \quad (33)$$

$$supply_{ik} \geq 0 \quad \forall i \in N, k = 1, \dots, K \quad (34)$$

$$flow_{ijk} \geq 0 \quad \forall (i, j) \in E, k = 1, \dots, K \quad (35)$$

The length of ST is computed in constraints (23). The following constraints are used to model ST and ensure the connectivity of the district. ST can only include edges that connect areas belonging to the district (constraints 24 and 25) and, therefore, only those edges can have a flow (26). ST can only have one root (27) which is chosen among the areas belonging to the district (28). Only the root can have a flow supply (29) which is set to the number of areas in the district (30). In the flow equilibrium constraints (31) only the areas belonging to the district require one unit of flow, thus, enforcing connectivity. Finally, the variables' existence conditions are set in constraints (32)-(35).

Acknowledgements The authors would like to thank the Spanish National Police Corps for the crime data used in the case study. Also, the authors thank Dr Elena Rosa Pérez of the Spanish National Statistics Institute and the Complutense University of Madrid for the assistance given in obtaining the 2011 census data.

Declaration

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Andrejevic M, Gates K (2014) Big data surveillance: introduction. *Surveill Soc* 12(2):185–196
- Bezanson J, Edelman A, Karpinski S, Shah VB (2017) Julia: a fresh approach to numerical computing. *SIAM Rev* 59(1):65–98
- Bodily SE (1978) Police sector design incorporating preferences of interest groups for equality and efficiency. *Manag Sci* 24(12):1301–1313
- Braga AA, Papachristos AV, Hureau DM (2014) The effects of hot spots policing on crime: an updated systematic review and meta-analysis. *Justice Q* 31(4):633–663
- Braga AA, Brunson RK, Drakulich KM (2019) Race, place, and effective policing. *Ann Rev Sociol* 45:535–555
- Camacho-Collados M, Liberatore F (2015) A decision support system for predictive police patrolling. *Decis Support Syst* 75:25–37
- Camacho-Collados M, Liberatore F, Angulo JM (2015) A multi-criteria police districting problem for the efficient and effective design of patrol sector. *Eur J Oper Res* 246(2):674–684
- Chen H, Cheng T, Wise S (2017) Developing an online cooperative police patrol routing strategy. *Comput Environ Urban Syst* 62:19–29

- Chen H, Cheng T, Ye X (2019) Designing efficient and balanced police patrol districts on an urban street network. *Int J Geogr Inf Sci* 33(2):269–290
- Cohen MA (2017) The social cost of a racially targeted police encounter 1. *J Benefit-Cost Anal* 8(3):369–384
- Dewinter M, Vandeviver C, Beken TV, Witlox F (2020) Analysing the police patrol routing problem: a review. *ISPRS Int J Geo-Inf* 9(3):157
- Dunning I, Huchette J, Lubin M (2017) Jump: a modeling language for mathematical optimization. *SIAM Rev* 59(2):295–320
- Duque JC, Church RL, Middleton RS (2011) The p-regions problem. *Geogr Anal* 43(1):104–126
- Enayati S, Ozaltin O. Y, Mayorga M. E (2020) Designing ambulance service districts under uncertainty. In: Rios-Mercado R.Z. (ed) *Optimal districting and territory design*. Springer, Cham, pp 153–170
- Fagan J (2017) Recent evidence and controversies in the new policing. *J Policy Anal Manag* 36(3):690–700
- Fu C, Wolpin KI (2018) Structural estimation of a becker-ehrlich equilibrium model of crime: allocating police across cities to reduce crime. *Rev Econ Stud* 85(4):2097–2138
- Galiani S, Cruz IL, Torrents G (2018) Stirring up a hornets' nest: geographic distribution of crime. *J Econ Behav Org* 152:17–35
- Glover F (1989) Tabu search-part I. *ORSA J Comput* 1(3):190–206
- Glover F (1990) Tabu search-part II. *ORSA J Comput* 2(1):4–32
- Gurobi Optimization L (2020) Gurobi optimizer reference manual Last accessed April 7, 2021. <http://www.gurobi.com>
- INE, Instituto Nacional de Estadística (Spanish National Statistics Institute) (2011) Censos 2011, Resultados (census 2011, results)
- Kim Y-A (2018) Examining the relationship between the structural characteristics of place and crime by imputing census block data in street segments: is the pain worth the gain? *J Quant Criminol* 34(1):67–110
- Kim H, Kim K (2020) Spatial optimization problem for locating polling facilities and stations and policy implications. In: Rios-Mercado R.Z. (ed) *Optimal districting and territory design*. Springer, Cham, pp 173–190
- Kleinberg J, Ludwig J, Mullainathan S, Rambachan A (2018) Algorithmic fairness. In: *Aea papers and proceedings* 108:22–27
- Leigh J, Dunnett S, Jackson L (2019) Predictive police patrolling to target hotspots and cover response demand. *Ann Oper Res* 283(1–2):395–410
- Liberatore F, Camacho-Collados M (2016) A comparison of local search methods for the multicriteria police districting problem on graph. *Math Probl Eng*. <https://doi.org/10.1155/2016/3690474>
- Liberatore F, Camacho-Collados M, Vitoriano B (2020) Police districting problem: literature review and annotated bibliography. In: Rios-Mercado R. Z. (ed) *Optimal districting and territory design*. Springer, Cham, pp 9–29
- Maheshri V, Mastrobuoni G (2019) The race between deterrence and displacement: theory and evidence from bank robberies. *Rev Econ Stat*. https://doi.org/10.1162/rest_a_00900
- Mitchell PS (1972) Optimal selection of police patrol beats. *J Crim L Criminol Police Sci* 63:577
- Mohler GO, Short MB, Malinowski S, Johnson M, Tita GE, Bertozzi AL, Brantingham PJ (2015) Randomized controlled field trials of predictive policing. *J Am Stat Assoc* 110(512):1399–1411
- Moya-García J.G., Salazar-Aguilar M. A. (2020) Territory design for sales force sizing. In: Rios-Mercado R.Z. (ed) *Optimal districting and territory design*. Springer, Cham, pp 191–2006
- National Academies of Sciences E, Medicine, et al., (2018) *Proactive policing: effects on crime and communities*. National Academies Press
- NOMECALLES, Nomenclátor oficial y callejero (Official gazetteer and street map), (2020) Selección de shapes a descargar (selection of shapes to download) Last accessed April 7, 2021. <https://www.madrid.org/nomecalles/DescargaBDTCorte.icm>
- OECD (2020) OECD Better Life Index <http://www.oecdbetterlifeindex.org/>
- Perry W (2013) L. The role of crime forecasting in law enforcement operations. Rand corporation, predictive policing
- Ricca F, Scozzari A (2020) Mathematical programming formulations for practical political districting. In: Rios-Mercado R.Z. (ed) *Optimal districting and territory design*. Springer, Cham, pp 105–128
- Rinehart Kochel T (2011) Constructing hot spots policing: unexamined consequences for disadvantaged populations and for police legitimacy. *Crim Justice Policy Rev* 22(3):350–374
- Ringuest J. L. (1992) *Compromise programming. Multiobjective optimization: behavioral and computational considerations*. Springer, New York, pp 51–59
- Rios-Mercado R. Z. (ed) (2020) *Optimal districting and territory design*. International series in operations research and management science. Springer, Cham

- Sampson RJ, Wilson WJ (1995) Toward a theory of race, crime and urban inequality. Stanford University Press, Stanford, CA, pp 37–56
- Scott MS, Clarke RV (2020) Problem-oriented Policing, Successful case studies. Routledge, New York
- Shirabe T (2005) A model of contiguity for spatial unit allocation. *Geogr Anal* 37(1):2–16
- Wheeler AP (2019a) Allocating police resources while limiting racial inequality. *Justice Q* 37:1–27
- Wheeler AP (2019b) Creating optimal patrol areas using the p-median model. *Int J Polic* 42(3):318–333
- Wikipedia contributors, (2020) George Floyd protests – Wikipedia, the free encyclopedia
- Yanık S, Bozkaya B (2020) A review of districting problems in health care. Optimal districting and territory design. Springer, Cham, pp 31–55

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Federico Liberatore^{1,2}  · **Miguel Camacho-Collados**^{3,4} · **Lara Quijano-Sánchez**^{2,5}

¹ School of Computer Science & Informatics, Cardiff University, Cardiff, United Kingdom

² UC3M-Santander Big Data Institute (IBiDat), Charles III University of Madrid, Madrid, Spain

³ Council of State, Madrid, Spain

⁴ Ministry of the Interior, Madrid, Spain

⁵ Engineering School, Autonomous University of Madrid, Madrid, Spain